

# Soil porous system as heterogeneous complex network

J.P. Cárdenas <sup>a</sup>, A. Santiago <sup>a</sup>, A.M. Tarquis <sup>b</sup>, J.C. Losada <sup>a</sup>, F. Borondo <sup>c</sup>, R.M. Benito <sup>a,\*</sup>

<sup>a</sup> Grupo de Sistemas Complejos, Departamento de Física, Universidad Politécnica de Madrid, 28040 Madrid, Spain

<sup>b</sup> Departamento de Matemática Aplicada, Universidad Politécnica de Madrid, 28040 Madrid, Spain

<sup>c</sup> Departamento de Química, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain

## ABSTRACT

In this paper we present a complex network model based on a heterogeneous preferential attachment scheme as a new way to quantify the porous structure of soils and to relate them with soil texture. We consider the pores as nodes, the properties of which, such as position and size, are described by fixed states in a metric space. An affinity function is introduced to bias the attachment probabilities of links according to these properties. We perform an analytical study of the degree distributions in the model and develop a numerical analysis of the degree distributions in the model for a combination of parameters corresponding to eleven empirical soil samples with different physical properties and five different textures.

### Keywords:

Complex networks

Porous structure

Dynamical model

## 1. Introduction

After soil texture, the geometry of pore space as well as pore size is probably one of the most important factors to understanding the transport of water, gas and solute in soils (Ball et al., 2007). However, a quantitative and explicit characterization, by means of a physical interpretation, is difficult because of the complexity of the pore space (VandenBygaert et al., 1999; Leij et al., 2002). This is the main reason why most theoretical approaches to soil porosity are idealizations to simplify the complex structure of the wiring and spatial location of pores (Bird et al., 2006). The complex interplay between the size and continuity of pores is considered to be the key in understanding the transport properties (Klute and Dirksen, 1986; VandenBygaert et al., 1999). Despite the empirical observation of the pore size distribution (PSD) and pore geometry distribution (PGD) (Lin et al., 1999; Vogel and Roth, 2001), the understanding of the emergence of these PSD and PGD is not entirely clear (Horgan and Ball, 1994). Considering this problem, in this work we propose a network model in order to analyze porous structures previously mentioned.

A network model is a flexible way of representing objects (nodes) and their relationships (links or connections) from a purely topological point of view showing it as a graph (Biggs et al., 1986). During the last decade, a wide range of real complex systems was studied from this perspective by means of Network Theory (Strogatz, 2001; Albert and Barabási, 2002). The results of these studies showed non-intuitive and non-trivial topologies in

the studied networks, irrespective of their origin. This means that the patterns of connection between their nodes are neither purely regular nor purely random, reason why they are called complex networks.

Many real complex networks are *scale-free* and display properties of *small-world* networks (Watts and Strogatz, 1998). A network is named *scale-free* if its degree distribution, i.e., the probability that a node selected uniformly at random has a certain number of links (degree), follows a power law implying that the distribution of connectivities of the network has no characteristic scale (Dorogovtsev and Mendes, 2002; Newman, 2003) denoting a high inhomogeneity in the connectivity of the system elements. In *scale-free* topologies some nodes, called "hubs", present a degree that is orders of magnitude larger than the average. These hubs are absent in random graphs where most nodes have a connectivity close to the average (Erdős and Rényi, 1959). On the other hand, a complex network is called *small-world* when it shows an average short path length between nodes and a high mean clustering, denoting respectively an efficient pattern of connectivity and high levels of transitivity in the network. The presence of these properties has created a large follow-up debate among physicists because of their ubiquity and seems to be a characteristic feature of critical-point behavior and a fingerprint of self-organized systems (Bak et al., 1987).

In our proposed model we interpret porous soils as heterogeneous networks where pores are represented by nodes with distinct properties (such as area and spatial location) and the links represent fluid flows between them. The networks of pores are generated by a model known as Heterogeneous Preferential Attachment (HPA) (Santiago and Benito, 2007a,b, 2008, 2009), a generalization of the Barabási-Albert (BA) model (Barabási and Albert, 1999) to heterogeneous networks. The BA model is based on the mechanisms of growth and preferential attachment (Price, 1965, 1976) and provides a minimal account of the process leading to the emergence of *scale-free* networks (Barabási et al.,

1999). In the BA model the process starts with a seed of arbitrary size and topology. A new node is added to the network at each step, bringing a fixed number  $m$  of links attached. These links are preferentially connected to the already existing nodes following the so-called *attachment rule*: the linking probability of a network node  $v_i$  is proportional to its degree  $k_i$ ,  $\Pi(v_i) = k_i / \sum_j k_j$ . This step is iterated until a desired number  $N$  of nodes have been added to the network. The HPA model incorporates the influence of node attributes in addition to the connectivity degree to the attachment rule.

Network models are prescriptive systems that generate networks with certain topological traits, such as degree or clustering distributions. Within these, dynamical network models (Dorogovtsev and Mendes, 2002), such as BA and HPA models, are stochastic discrete-time dynamical systems that evolve networks by the iterated addition and subtraction of nodes and links. Until recently, soil pore network has been a detailed model of a porous medium, generally incorporating pore-scale descriptions of the medium and the physics of pore-scale events. It has been used to describe a wide range of properties from capillary pressure characteristics to interfacial area and mass transfer coefficients. The pores and throats are assigned some idealized geometry and rules are developed to determine the multiphase fluid configurations and transport in these elements (Berkowitz and Ewing, 1998) based on a random or lattice structure. Recently, taking into account the topology of real complex systems, two works have introduced a pore network concept from a topological point of view (Santiago et al., 2008; Mooney and Korošak, 2009) opening a new line in soil science context.

Under this new approach to the complex structure of the porous system we present this work as follows. In the next section we motivate and describe the formulation of the porous soil model based on the HPA model. In the third section we present the analytical solution for the degree distribution in the pore networks generated by the HPA model. In the fourth section we present the numerical results of the model corresponding to simulations of eleven real samples of soil with different structure and physical properties. In the last section we present our conclusions derived from this work.

## 2. Model formulation

The structure of a porous soil is modeled as a heterogeneous complex network where nodes  $v_i$  correspond to pores and links  $e_{ij}$  correspond to fluid flows between them. Fig. 1 briefly illustrates this idea. The links will

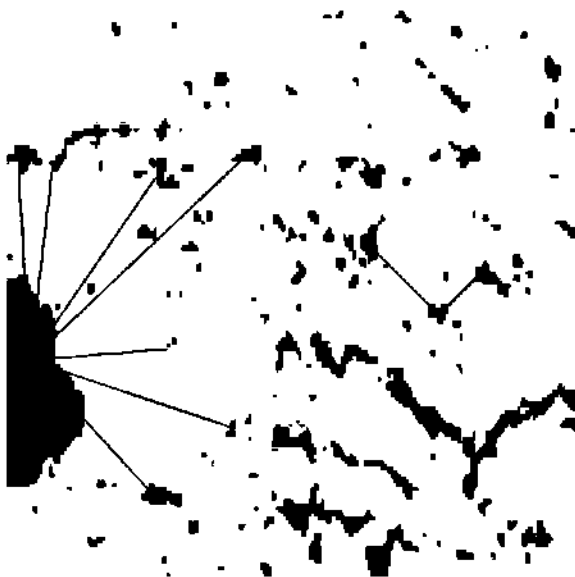


Fig. 1. Cartoon of a soil pore network. Red lines correspond to links between pores.

be considered undirected  $e_{ij} = e_{ji}$  and thus the connectivity degree  $k_i$  of a node  $v_i$  will be a measure of the number of pores directly connected with the associated pore. The properties of a pore are described by the node state  $(r_i, s_i)$ , which account for the position  $r_i$  of the pore center in the soil and the pore size  $s_i$ .

The evolution of the porous soil structure is modeled as a stochastic growth process described by a dynamic network model. Assuming that at a given time a new pore is created, the likelihood that it will connect with any of the already existing pores will be proportional to the size of the older pores and inversely proportional to the distance between them. Likewise, the higher the number of connections accruing to an existing pore, the higher the likelihood that a new flow of fluids will intercept an existing connection and connect the new pore with the older one. Thus the attachment visibility of an existing network node  $\pi(v_i)$  when a new node  $v_a$  is added will be a monotonous function of  $k_i, s_i$  and  $d^{-1}(r_i, r_a)$ , where  $d$  is the Euclidean metric.

The previous considerations prompt us to model the dynamics of the porous structure of soils as a particular case of HPA model defined by a state space  $R$ , a probability distribution  $\rho(x)$  of the node states  $x \in R$ , and an affinity function,  $\sigma(x, y)$ , of the interactions between the existing and added nodes that depends on their states. This formalism prescribes the evolution of a network according to the following rules:

- (i) The nodes  $v_i$  are characterized by their state  $x_i \in R$  deemed constant in the timescale of evolution of the network.
- (ii) The growth process starts with a seed composed by  $N_0$  nodes (with arbitrary states  $x_i \in R$ ) and  $L_0$  links.
- (iii) A new node  $v_a$  (with a fixed number  $m$  of links attached) is added to the network at each iteration. The newly added node is randomly assigned a state  $x_a$  following the distribution  $\rho(x)$ .
- (iv) The  $m$  links attached to  $v_a$  are randomly connected to the network nodes following a distribution  $\{\Pi(v_i)\}$  given by an extended *attachment rule*,

$$\Pi(v_i) = \frac{\pi(v_i)}{\sum_j \pi(v_j)}, \quad \pi(v_i) = k_i \cdot \sigma(x_i, x_a). \quad (1)$$

The *attachment kernel* or *visibility*  $\pi$  of a node  $v_i$  in the rule is given by the product of its degree  $k_i$  and its affinity  $\sigma$  with the newly added node  $v_a$ , which is itself a function of the states  $x_i$  and  $x_a$ . It thus can be seen that for each interaction  $\sigma$  biases the degree  $k_i$  of the candidate node. Steps (iii) and (iv) are iterated until a certain number of nodes have been added to the network. This number is determinate by the number of soil pores estimated in the 2D image. To sum up, the choice of the triple  $(R, \rho, \sigma)$  determines the form of heterogeneity in the attachment mechanism.

Consistently with the previous formalism, the porous soil model is defined by a state space  $R = M \times S$ , where  $M$  is a Euclidean box with dimension 2 or 3 that represents the medium geometry and  $S$  is an interval of the real line that represents the spectrum of possible pore sizes; a state distribution  $\rho(x) = \rho(r, s)$  that represents the probability for a new pore having a certain position  $r$  and size  $s$ ; and an affinity function

$$\sigma(x, y) = \frac{s_x^\alpha}{(\delta + |r_x - r_y|)^\beta}, \quad (2)$$

where  $\delta$  is a small nonnegative offset that ensures the continuity of  $\sigma$  on  $R^2$ , while  $\alpha$  and  $\beta$  are free parameters that measure the relative importance of the size and distance of the pores in the affinity of the attachment mechanism. Defining the porosity of the medium  $\phi = S_v / S_T = \sum s_i / \int_M dr$  as the fraction of void space in the material, then the desired number of nodes  $N$  that ends the growth process is the least number that yields the porosity of the simulated soil  $\phi > \phi_0$  for a certain threshold  $\phi_0$ , which is the porosity of the sample medium.

### 3. Analysis

In this section we present an analytical study of the stationary degree distribution  $P(k)$  of the proposed soil model. The solution is obtained by rate equations (Dorogovtsev et al., 2000; Krapivsky and Redner, 2001) which establish a balance in the connections degree densities over a partition of the network. The exposition will be brief therefore we submit the reader to Santiago et al. (2008) for a more detailed discussion of the solution, and to Santiago and Benito (2008) for a discussion of the solution for the general class of preferential attachment (PA) models. Let us first define a sequence of functions  $\{f(k, x, N)\}_{N \geq 0}$  which measure the probability density of a randomly chosen node having degree  $k$  and state  $x$  in a network at the iteration  $t = N$ . The degree densities are local metrics, thus they uniformly converge when  $N \rightarrow \infty$  to a stationary density function  $f(k, x)$ . Finally, the stationary degree distribution  $P(k)$  measures the probability of a randomly chosen node having degree  $k$  in the thermodynamic limit.

We will denote by  $V(x) = \{v_i, x_i = x\}$  the subset of nodes in the network with state  $x = (x_i, s_i)$ . Assuming that the assignation of states  $x_i$  is uncorrelated with the topology of the growing network, and that there are no linking events between existing nodes, the sequence  $\{f\}$  can be modeled on each  $V(x)$ . For each  $x$ , the form of the equation will be  $L_1 - L_2 = R_1 - R_2$ , where:

- $L_1$  density of nodes with degree  $k$  at  $t = N + 1$ ;
- $L_2$  density of nodes with degree  $k$  at  $t = N$ ;
- $R_1$  increase in density due to nodes with degree  $k - 1$  that have gained a link at  $t = N$ ;
- $R_2$  decrease in density due to nodes with degree  $k$  that have gained a link at  $t = N$ .

The resulting density rate equation for  $k > m$  is

$$(N + 1)f(k, x, N + 1) - Nf(k, x, N) = m \left\langle \frac{\sigma(x, y)}{\psi(y, N)} \right\rangle_y [(k - 1)f(k - 1, x, N) - kf(k, x, N)], \quad (3)$$

where  $\psi(y, N)$  is defined as the partition factor

$$\Psi(y, N) = \sum_k k' \int_R \sigma(x, y) f(k', x, N) dx, \quad (4)$$

and the brackets mean averaging over the random variable  $y$ ,

$$\langle g \rangle_y = \int_R g p(y) dy. \quad (5)$$

For  $k = m$  the resulting density rate equation is

$$(N + 1)f(m, x, N + 1) - Nf(m, x, N) = \rho(x) - m \left\langle \frac{\sigma(x, y)}{\psi(y, N)} \right\rangle_y mf(m, x, N). \quad (6)$$

There are no nodes with degree  $k < m$ , since when  $N \rightarrow \infty$  all the links attached to newly added nodes find receptive nodes in the network, therefore the previous equations define all the possible cases in each iteration.

In the thermodynamic limit  $N \rightarrow \infty$ ,  $f(k, x, N + 1) = f(k, x, N) = f(k, x)$  and the rate equations become

$$\begin{aligned} f(k, x) &= \\ m \left\langle \frac{\sigma(x, y)}{\psi(y)} \right\rangle_y [(k - 1)f(k - 1, x) - &kf(k, x)] \quad \text{for } k > m \\ \rho(x) - m \left\langle \frac{\sigma(x, y)}{\psi(y)} \right\rangle_y mf(k, x) &\quad \text{for } k = m \end{aligned} \quad (7)$$

where  $\psi(y)$  is the stationary partition factor, defined as

$$\begin{aligned} \Psi(y) &= \lim_{N \rightarrow \infty} \Psi(y, N) = \\ &= \sum_{k'} k' \int_R \sigma(x, y) f(k', x) dx. \end{aligned} \quad (8)$$

Through a mean-field approach it is possible to decouple the previous equations under the approximation criteria

$$\langle (\Psi(y) - \psi)^2 \rangle_y < \varepsilon_1, \quad \langle (w(x) - \bar{w})^2 \rangle_x < \varepsilon_2, \quad (9)$$

where  $\psi = \langle \Psi(y) \rangle_y$  is the mean-field partition factor of the network,  $w(x) = \langle \sigma(x, y) \rangle_y$  is the mean-field fitness of network nodes with state  $x$ ,  $\bar{w} = \langle w(x) \rangle_x$  is the network average mean-field fitness, and  $\varepsilon_1$  and  $\varepsilon_2$  are accuracy bounds. This renders the decoupled system

$$\begin{aligned} f(k, x) &= \\ \hat{w}(x) [(k - 1)f(k - 1, x) - kf(k, x)] / 2 &\quad \text{for } k > m, \\ \rho(x) - \hat{w}(x) mf(k, x) / 2 &\quad \text{for } k = m. \end{aligned} \quad (10)$$

where  $\hat{w}(x) = w(x) / \bar{w}$  is a mean-field normalized fitness, which measures the fitness of network nodes with state  $x$  relative to the average fitness of all the network nodes, for a given distribution  $\rho(x)$ .

For the sake of brevity, in the foregoing the dependence of functions  $\rho(x)$  and  $\hat{w}(x)$  on the state  $x$  will no longer be explicated, i.e.  $\rho = \rho(x)$  and  $\hat{w} = \hat{w}(x)$ . Solving Eq. (10) we obtain  $f(m, x) = 2\rho / (\hat{w}m + 2)$  and the solution for the stationary density for  $k > m$  is

$$f(k, x) = \left( \prod_{j=m+1}^k \frac{\hat{w}(j-1)}{\hat{w}j + 2} \right) \frac{2\rho}{\hat{w}m + 2} \quad (11)$$

and integrating the density in Eq. (11) over the state space  $R$  the stationary degree distribution is

$$P(k) = \int_R \left( \prod_{j=m+1}^k \frac{\hat{w}(j-1)}{\hat{w}j + 2} \right) \frac{2\rho}{\hat{w}m + 2} dx. \quad (12)$$

The solution given by Eq. (12) is valid for any variant of the soil model, irrespective of the geometry of the medium  $M$  or the pore sizes  $S$ , the distribution of pore properties  $\rho$  or the dependence of the affinity  $\sigma$  on the pore properties via  $\alpha$  and  $\beta$ . Nevertheless, the validity of the solution is restricted by the extent to which the approximation criteria in Eq. (9) hold. The mean-field approximation can be considered accurate in the measure that the averages of  $\sigma(x, y)$  along its two arguments do not exhibit large variations over  $R$ .

Furthermore, given that any pore in the soil has a non-zero size, it follows that by definition the mean-field fitness  $\hat{w}(x)$  is strictly

**Table 1**  
Classification, horizon, depth and physical composition of soil samples.

Soil	Classification	Horizon	Depth	Org. mat.	Sand	Silt	Clay
Inter	Argiaquic Argialboll	B	34–40	0.4	3.1	66.2	25.1
Kampong	Typic Haplorthox	B	34–38	1.2	20.0	22.0	58.0
Munchong	Typic Haplorthox	B	70–78	0.7	22.0	6.0	72.0
Site	Inceptisol	C	20–25	0.3	94.6	4.9	0.4
Buso	Spodosol	C	75–63	0.1	54.0	15.0	31.0
Evhl	Entisol	A	1–10	0.8	37.0	30.5	32.5
Abok	Orthoxic Tropudult	B	33–40	0.3	48.0	21.0	31.0
Denny	Mollic Albaqualf	B	47–52	0.4	2.7	67.1	30.2
Durian	Petric Plinthudult	A	1–8	1.7	12.0	51.0	37.0
Mss	Entisol	BC	20–32	1.3	2.7	63.5	33.8
Ads	Entisol	B	6–14	1.1	1.4	65.2	33.4

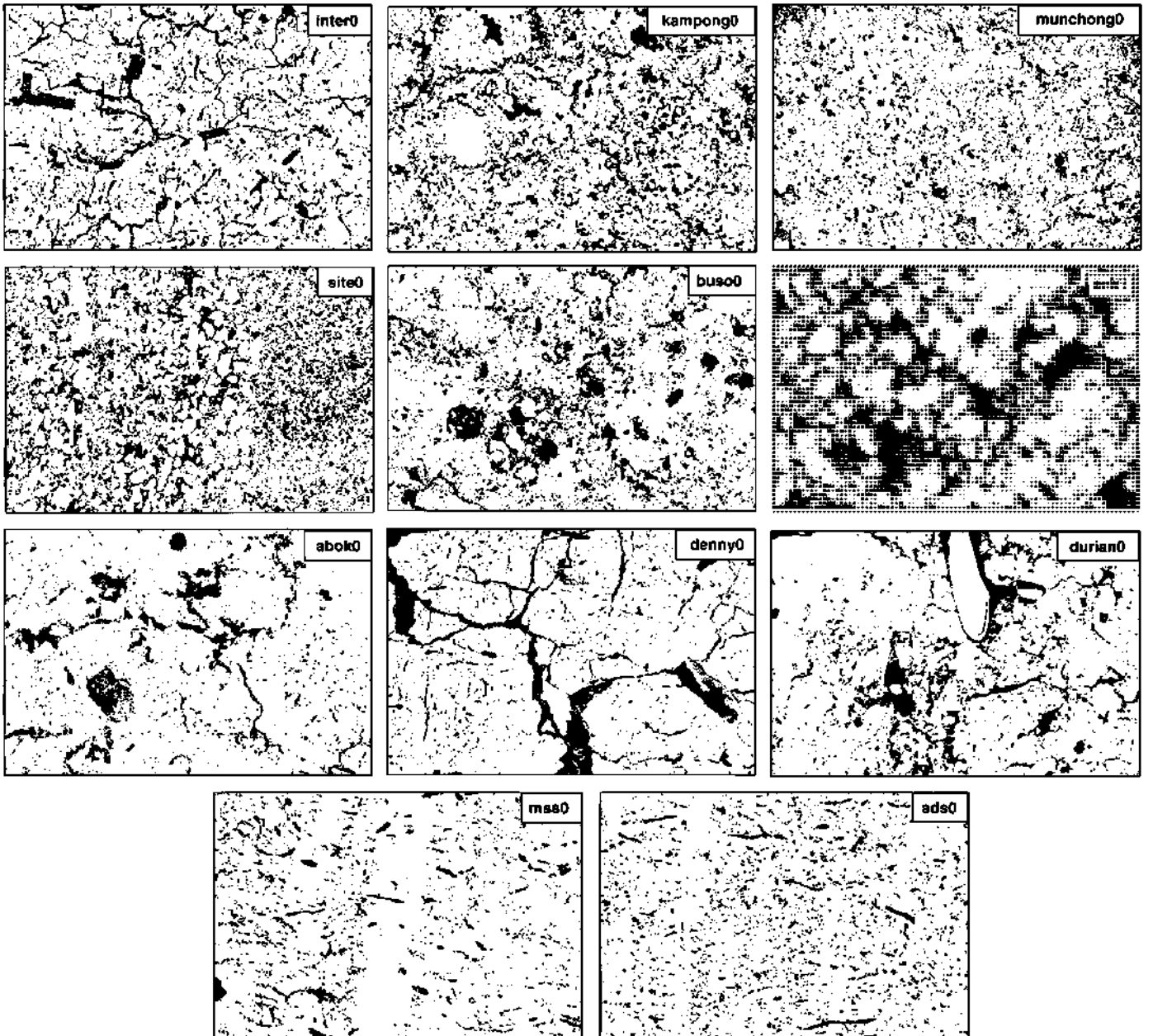
positive over  $R$  and the stationary degree density in Eq. (11) may be expressed for  $k \geq m$  as

$$f(k, x) = \frac{2\rho/\hat{\omega}}{m + 2/\hat{\omega}} \frac{B(k, 1 + 2/\hat{\omega})}{B(m, 1 + 2/\hat{\omega})}, \quad (13)$$

where Legendre's Beta function  $B(y, z) = \int_0^1 t^{y-1}(1-t)^{z-1}dt$  for  $y, z > 0$  satisfies the functional relation  $\Gamma(a)/\Gamma(a+b) = B(a, b)/\Gamma(b)$  for Euler's Gamma function  $\Gamma(x)$ . Likewise, integrating the density in Eq. (13) over  $R$  we obtain an expression for the stationary degree distribution  $P(k)$  for  $k \geq m$  in terms of Legendre's Beta function which is equivalent to Eq. (12).

Notice that the stationary densities for a given pore state, as given by Eq. (13), follow a Beta function with arguments  $(k, 1 + 2/\hat{\omega})$ . This implies that the degree densities exhibit a *multiscaling* according to power laws  $k^{-\gamma(\hat{\omega})}$  along the continuous spectrum of normalized

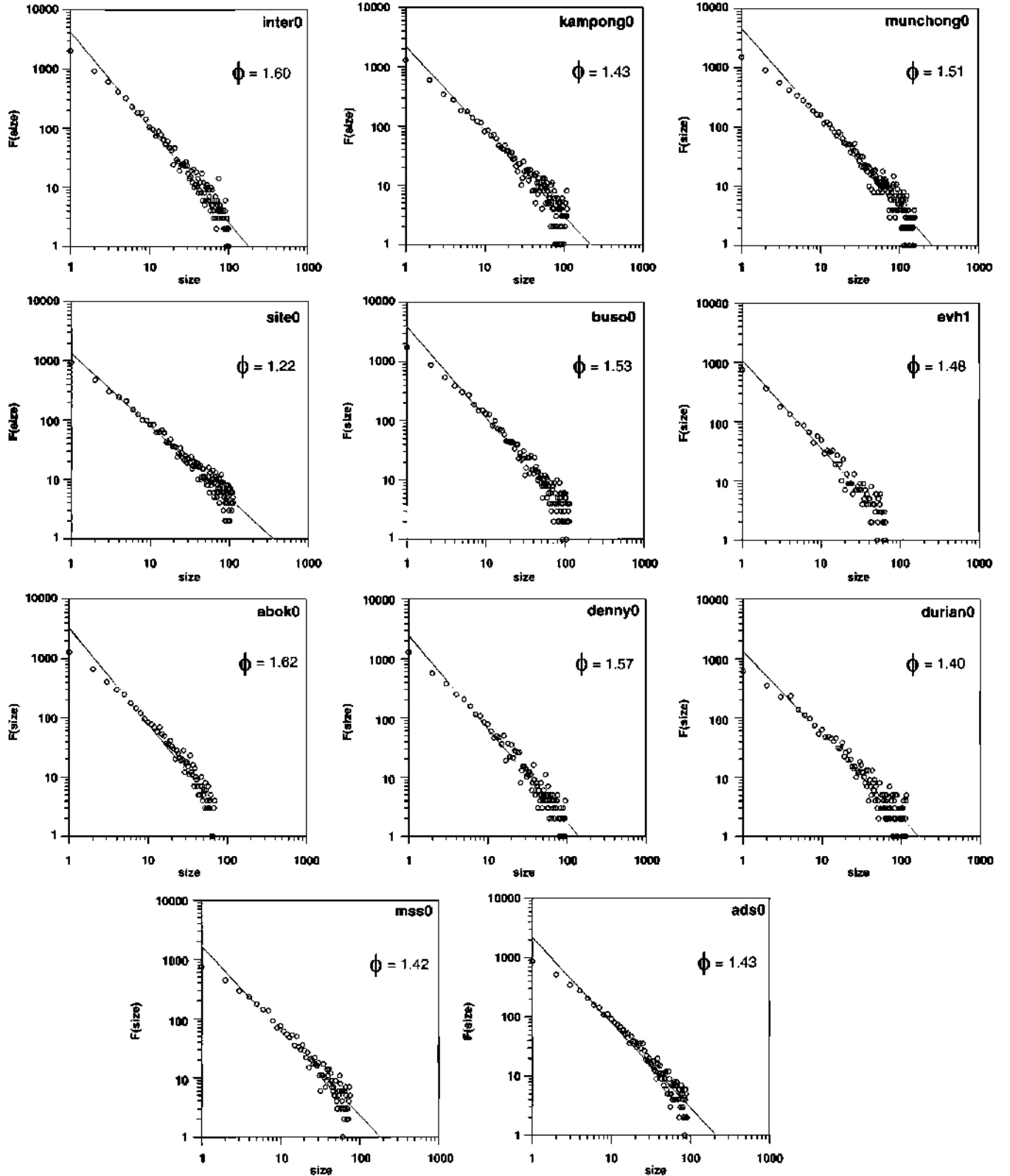
fitness  $\hat{\omega}$ , with scaling exponents  $\gamma(\hat{\omega}) = 1 + 2/\hat{\omega}$  spanning themselves a continuum. The multiscaling phenomenon is a general property of heterogeneous PA networks (Santiago and Benito, 2007a, 2008) that is exhibited by any variant of the proposed soil model, irrespective of the particular details. As  $\hat{\omega}$  increases (resp. decreases), the exponent  $\gamma$  of  $f$  decreases (resp. increases). Nodes with states more fit than the average ( $\hat{\omega} > 1$ ) adopt densities with  $\gamma < 3$ , which exhibit a slower asymptotical decay and tend to produce more hubs. The increase of either  $\alpha$  or  $\beta$  results in an increase in the variability of  $\omega$  over  $R$ , and in the case of  $\beta$  such an increase is more acute as the inhomogeneity of  $\rho_M$  increases. The higher variability of  $\omega$  translates into a larger spread of the distribution of exponent  $\gamma$  of the density components. This evidences a *signature* of heterogeneous PA in the structure of porous soils, by which the density components  $f(k, x)$  of pores will exhibit different scaling exponents according to their intrinsic properties (Santiago et al., 2008).



**Fig. 2.** Binary images of soil samples used in the study. Image size: 1500 × 1000 pixels.  
Credit: Richard Heck.

Eq. (12) shows that the degree distribution  $P(k)$  is obtained by the integration of power laws with varying exponents  $\gamma(x)$ . As we have seen, little fluctuations of  $\omega$  over  $R$  will yield density components with similar scaling exponents and thus distributions similar to the homogeneous PA case. On the other hand, large fluctuations of  $\omega$

over  $R$  (as in the case of large  $\alpha$  or large  $\beta$  parameters) will yield a wider spectrum of  $\gamma$  and thus a larger deviation of  $P(k)$  from the BA model. The asymptotic behavior of  $P(k)$  will be dominated by the slowest decaying components, associated to properties with highest  $\omega$ . Furthermore, given that  $\hat{\omega}$  is distributed by definition around 1, the



**Fig. 3.** Pore size distribution,  $F(\text{size})$ , for soil samples. The value of  $\phi$  corresponds to the scaling exponent of the fit to a power law denoted by the straight lines. Pore size is expressed as number of pixels.

highest value of the spectrum will verify  $\phi_{\max} \geq 1$  and therefore the degree distributions  $P(k)$  of all the soil variants will exhibit exponents satisfying  $1 < \gamma \leq 3$ . Finally, it is worth noting that as the sample size becomes lower the finite-size effects will become more important, yielding a growing discrepancy with the analytical results, such as a decrease in the scaling exponent  $\gamma$  of the degree distribution and the presence of an exponential cutoff regime. This problem could nevertheless be circumvented by modeling a larger soil sample. In our case, size sample and resolution of the images were enough for the porpoise of our study.

#### 4. Model application

In this section we present results concerning the numerical simulation of the porous soil model using as input data from eleven intact soil samples with different physical properties. We simulate pore networks with the same porosity, size distribution and spatial location of the empirical 2D soil samples. The soil samples were primarily selected to include contrasting soil structures and textures (Table 1). Each of the samples was prepared for image analysis following the procedure described by Protz and VandenBygaart (1998). These sample soils were kindly provided by University of Guelph, Department of Land Resource Science. For further information about image acquisition and soil properties we refer to the Canadian Soil Thin Section Collection website at URL: <http://gis.lrs.uoguelph.ca/cstsc/> focused on sections prepared for the 1978 ICSS in Edmonton. The data were obtained by imaging thin sections with a Kodak 460 RGB camera using transmitted and circularly polarized illumination. The data were cropped from  $3060 \times 2036$  pixels to  $3000 \times 2000$  pixels. Then EASI/PACE software classified the data and the void bitmap separated (individual pixel size was  $18.6 \times 18.6 \mu\text{m}$ ).

The samples showed a wide range for the percentages of sand (1.4–94.6%), silt (4.9–67.1%), and clay (0.4–72.0%). Organic matter content was relatively low (0.1–1.7%) because mainly subsurface horizons were sampled. Binary images of the soil samples are shown in Fig. 2.

Due to the fact that our model considers the pore size (surface =  $s$ ) we analyze the distribution of pore sizes in the soil samples. The PSD for the samples, given by  $F(\text{size})$ , is presented in Fig. 3.

As can be seen, all the soils display a distribution that follows a power law denoting a *scale-free* character in the distribution of pore sizes,  $F(\text{size}) \sim \text{size}^{-\phi}$ , except in the extremes values. We could divide size axis in two or three ranges to better fit in each one the scaling exponent, nevertheless we prefer to follow previous works (Mooney and Korošak, 2009) and keep an average along all the sizes. In Table 2 we show the value of  $\phi$  obtained from the  $F(\text{size})$  distribution, as well as the number of pores, mean size of pores (pixels<sup>2</sup>) and maximum size of pores (pixels<sup>2</sup>) obtained from each sample. It is remarkable that samples with clearly different pore size distributions and with distinctive size ranges show very close scaling exponent value when

**Table 2**  
Scaling exponent  $\phi$  of  $F(\text{size})$ , number of pores, mean size of pores and maximum size of pores corresponding to soil samples.

Soil	$\phi$	Number of pores	Mean size	Max size
Inter	1.60	6435	8.78	100
Kampong	1.43	4634	12.12	110
Munchong	1.51	7102	14.80	155
Site	1.22	4273	16.09	112
Buso	1.53	6431	10.93	114
Evh1	1.48	2272	7.46	63
Abok	1.62	4541	7.95	70
Denny	1.57	4120	8.65	96
Durian	1.40	2847	12.92	116
Mss	1.42	3347	10.28	76
Ads	1.43	4113	11.68	90

**Table 3**

Texture of the soil samples and the corresponding parameters  $\alpha$  and  $\beta$  used in the model implementation.

Soil	Texture	$\alpha$	$\beta$
Inter	Silt loam	0.5	2.0
Kampong	Clay	1.0	2.0
Munchong	Clay	1.0	2.0
Site	Sand	0.5	1.0
Buso	Sandy clay loam	0.5	1.0
Evh1	Sandy clay loam	0.5	1.0
Abok	Sandy clay loam	0.5	1.0
Denny	Silt clay loam	0.5	2.0
Durian	Silt clay loam	0.5	2.0
Mss	Silt clay loam	0.5	2.0
Ads	Silt clay loam	0.5	2.0

all the sizes are considered to fit a power law. These values were used as data for the model implementation.

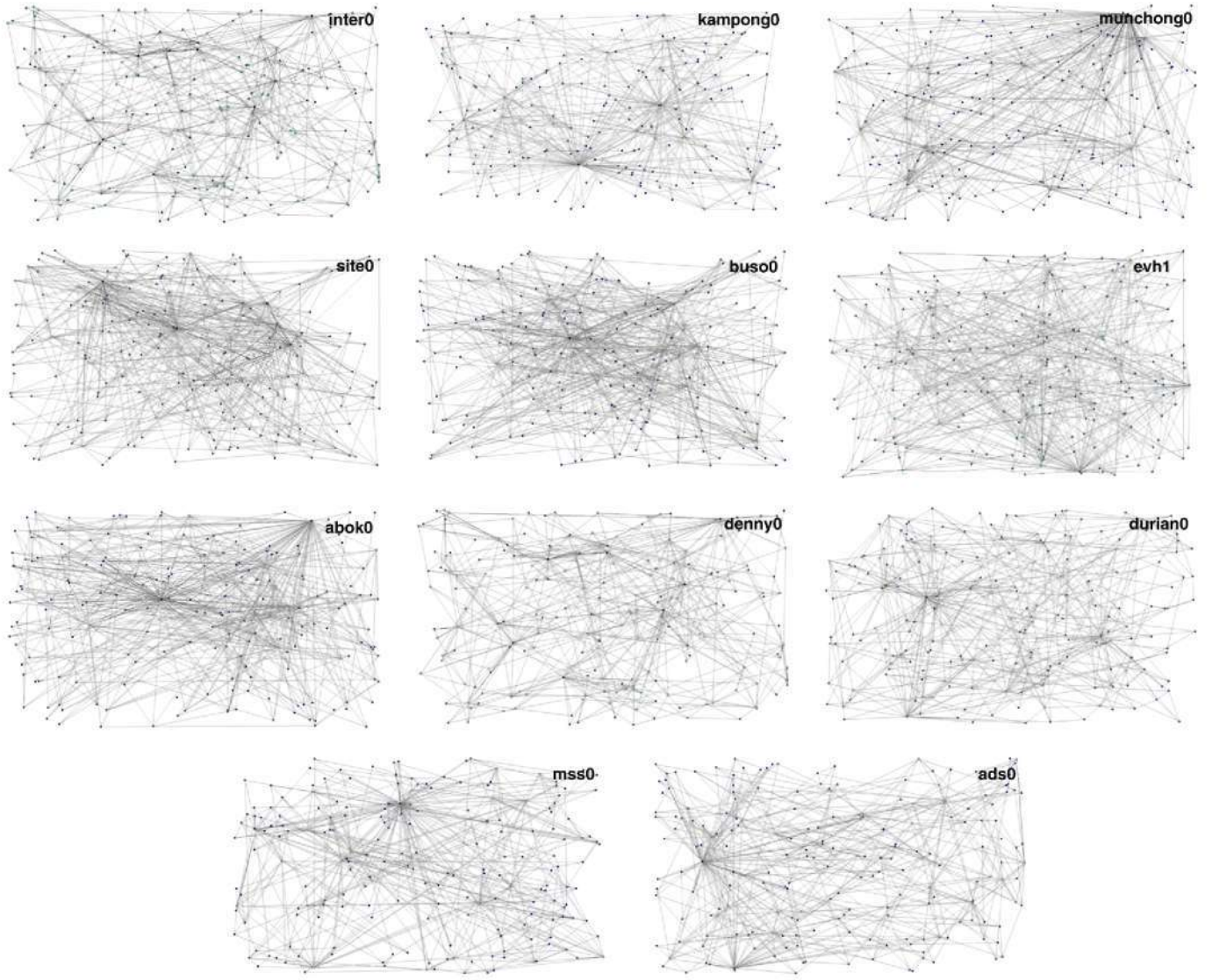
The affinity parameters of the model,  $\alpha$  and  $\beta$ , were chosen differently according to the textures of soils. We assume that soils with high percentages of clay and small size of pores are more compact in comparison with those composed by a high percentage of sand. For this reason, for soils with high percentage of clay we assign to the parameter  $\beta$ , associated to the distance between pores, a higher importance in our model. On the other hand, we also assume that the edge of a pore in our 2D simulation can be occupied by a higher number of particles in the case of small particles such as clay or silt in comparison to a pore of similar size but generated by large particles such as sand. Thus the number of potential connections (i.e., space that separates two particles) of a “clay pore” is higher in comparison with a “sandy pore”. For this reason, the parameter  $\alpha$ , associated to the surface of the pore, is less important in sandy soils. The values of  $\alpha$  and  $\beta$  used in the implementation for each soil texture are shown in Table 3.

#### 4.1. Results of the model

In the model all the generated networks start with a seed of  $N_0 = 10$  pores connected by  $L_0 = 9$  links. In each case the number of potential links for each new node added is  $m = 3$  and it remains unaltered during the evolution of the network. The aggregation process is iterated until 200 nodes have been added to the network (the low number is chosen for the sake of visibility). The Cartesian pore networks generated by the model are shown in Fig. 4. The networks layouts evidence that the size of the pores tends to be a more significant trait in the final connectivity of the pore than the pore position, as evidenced by a large share of highly connected nodes being located in peripheral positions of the soil surface. This insight may be explained by the fact that pore sizes follow highly inhomogeneous distributions (Perrier et al., 1996), while pore positions follow completely homogeneous distributions.

There is a remarkable presence of highly connected nodes in networks corresponding to soil with clay and sandy clay loam textures. In the case of the silt clay loam texture, the generated networks are characterized by the absence of highly connected pores. Looking with more attention to the soil images corresponding to this texture type (see Fig. 2) a common pattern is observed: longitudinal pores, which in the case of denny soil could be named cracks, are frequently found and between them very small pores are located. Even though links between the nodes in these complex networks do not reflect the true physical connections between pores, they characterize soil structure (Mooney and Korošak, 2009). Therefore,  $\alpha$  and  $\beta$  parameters are useful to relate network links with soil texture which have a huge influence in soil porosity and connectivity. Here we must emphasize that the organic matter content of the soil samples selected in this study is low (less than 1.8% as can be seen in Table 1).





**Fig. 4.** Cartesian networks of pores generated by the soil model considering real data from soil samples. Node position reflects the spatial location of a pore center of mass in the simulated soil sample. Node color corresponds to the connectivity degree  $k$ : warm colors mean high degrees, cold colors mean low degrees.

Fig. 5 depicts the degree distributions  $N \times P(k)$  obtained through numerical simulation of the generated networks with a size of  $N = 10,000$  pores. The scaling exponent  $\gamma$  of the degree distribution varies from 2.07 (munching sample) to 2.65 (mss sample) not showing a high numerical difference, however we should have in consideration that it is the exponent of a power law. In all the cases where  $a = 0.5$  and  $b = 2.0$  the degree distribution presents a  $k$  maximum range around 1000 (see Fig. 5). In the cases where  $b$  was reduced to 1 this maximum did not exist except for ehv1 that presents the lower number of pores and shorter range of pore size than the rest with these parameters (see Table 1). Finally, those cases where  $a = 1$ , kampong and munching, show a strong nonlinear trend in the degree distribution at the less connected pores. Even though these results do not reflect the variety of images shown in Fig. 2, they offer us information about the topology of these pore networks with more common features than expected.

The degree distributions do not fit well a power law along the studied interval, however they evidence a progressively better agreement for higher degrees. This can be explained by the fact that the power law components in the density spectrum with lower scaling exponents dominate the distribution decay, so that the asymptotic behavior of  $N \times P(k)$  over arbitrarily high degrees would

fit the power law as analytically predicted. The behavior of the distribution for lower degrees can be explained by the inhomogeneity of the distribution of pore sizes, which yields a non-negligible presence of very large pores in the network. We have to remark that the connectivity pattern between pores was not observed empirically, the *scale-free* distributions obtained (both numerically and analytically) are the result of the implementation of a reasonable attachment rule  $\Pi(v_i)$  and the adoption of real parameters of soil samples.

Finally, in this work the general scheme established a biasing dependence on node size and distance to be parameterized by power laws, and the dependence on degree being strictly linear. Future work should consider extending the formalism so as to include nonlinear functions of the degree.

## 5. Conclusions

In summary, we have presented a complex network model based on a heterogeneous preferential attachment scheme in order to quantify the structure of porous soils. The proposed network model allows the specification of different medium geometries and pore dynamics through the choice of different underlying spaces, distributions of

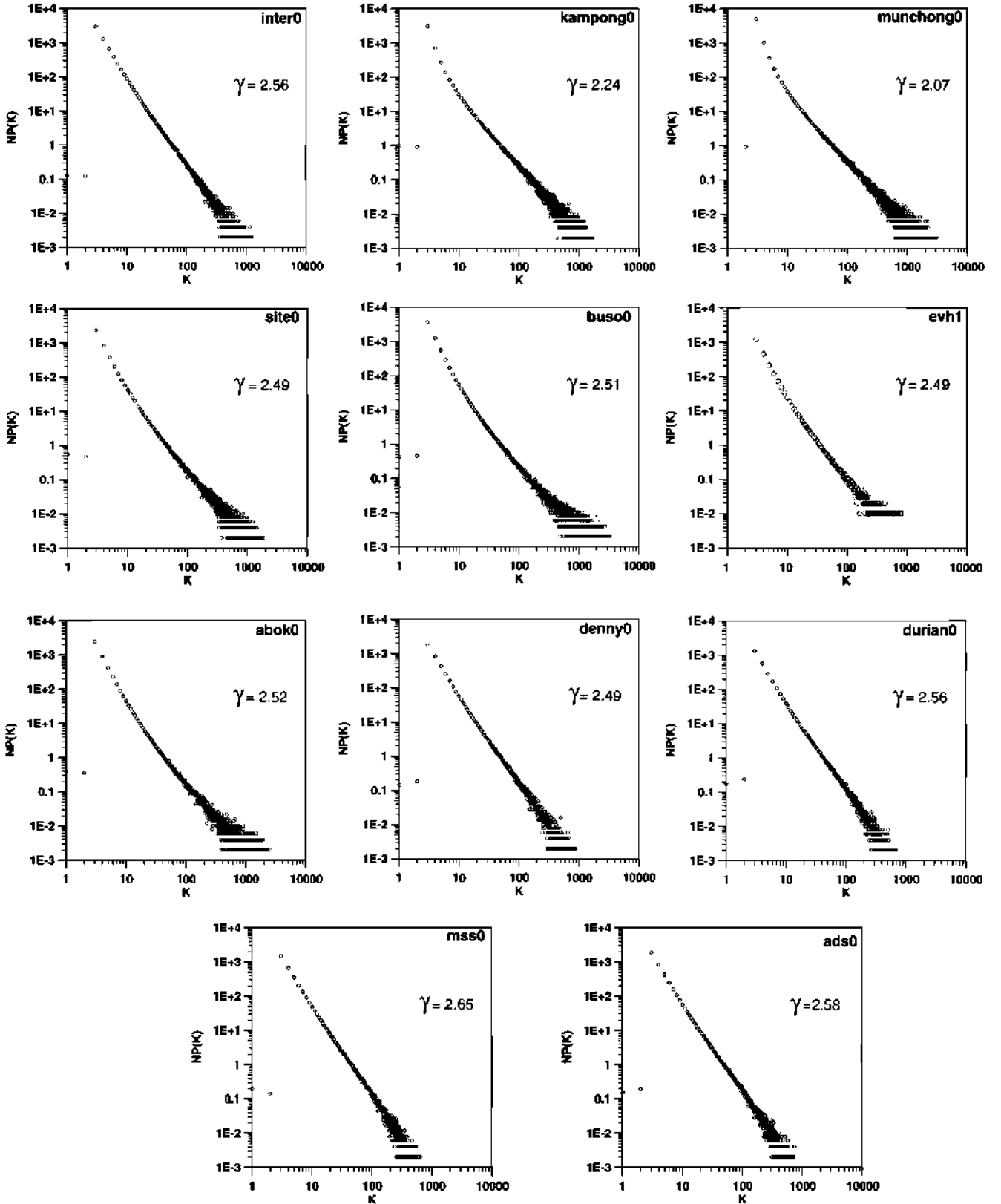


Fig. 5. Distribution of the number of nodes with degree  $k$ ,  $N \times P(k)$ , for the soil samples. The value of  $\gamma$  corresponds to the scaling exponent of the fit to a power law.

pore properties and affinity functions. We have obtained analytical solutions for the degree densities and degree distribution of the pore networks generated by the model in the thermodynamic limit. We have shown that these networks exhibit a multiscaling of their degree densities according to power laws with exponents spanning a

continuum, and that such phenomenon leaves a signature of heterogeneity in the topology of pore networks that can be empirically tested in real soil samples.

We have also shown the relationship between the variability in the scaling exponents and the parameters regulating the affinity function, as



well as the inhomogeneity of the distributions of pore properties, and the consequences of this on the asymptotic behavior of the degree distribution. In particular, it is worth emphasizing that the degree distributions of all the variants of the model exhibit power law behaviors with exponents within the limits empirically observed in real networks. We have performed a numerical analysis of the model for a combination of parameters corresponding to empirical samples with different properties. We have shown that the simulation results exhibit a good agreement with the analytical predictions, concerning the scaling behavior of the degree metrics as well as the ranges of values of the scaling exponents obtained.

Finally, it is worth remarking that the application of heterogeneous complex networks to the modeling of porous soils offers the possibility of linking microscopic traits (such as biasing parameters related to the soil texture) with non-trivial macroscopic properties (as illustrated by scale-free degree distributions and small-world properties), as well as furthering our understanding of the dynamics of soil development.

## Acknowledgements

This work has been supported by the Spanish MEC under Project 'i-MATH' No. CSD2006-00032 and Project No. MTM2009-14621, GESAN and Comunidad de Madrid under Project TAGRALIA-CM-P-AGR-000187-0505. We are grateful to Prof. Richard Heck, from the Dept. of Soils and Landscape Processes of Guelph University, for the images provided for this work.

## References

- Albert, R., Barabási, A.-L., 2002. Statistical mechanics of complex networks. *Rev. Mod. Phys.* 74 (1), 47–97.
- Bak, P., Tang, C., Wiesenfeld, K., 1987. Self-organized criticality: an explanation of  $1/f$  noise. *Phys. Rev. Lett.* 59, 381–384.
- Ball, B.C., Watson, C.A., Baddeley, J.A., 2007. Soil physical fertility, soil structure and rooting conditions after ploughing organically managed grass/clover swards. *Soil Use Manage.* 23, 20–27.
- Barabási, A.-L., Albert, R., 1999. Emergence of scaling in random networks. *Science* 286, 509–512.
- Barabási, A.-L., Albert, R., Jeong, H., 1999. Mean-field theory for scale-free random networks. *Physica A* 272, 173–197.
- Berkowitz, B., Ewing, R.P., 1998. Percolation theory and network modeling applications in Soil Physics. *Surv. Geophys.* 19, 23–72.
- Biggs, N., Lloyd, E., Wilson, R., 1986. *Graph Theory*. Oxford University Press.
- Bird, N., Cruz Díaz, M., Saab, A., Tarquis, A.M., 2006. Fractal and multifractal analysis of pore-scale images of soil. *J. Hydrol.* 322, 211–219.
- Dorogovtsev, S.N., Mendes, J.F.F., 2002. Evolution of networks. *Adv. Phys.* 51, 1079–1187.
- Dorogovtsev, S.N., Mendes, J.F.F., Samukhin, A.N., 2000. Structure of growing networks with preferential linking. *Phys. Rev. Lett.* 85, 4633–4636.
- Erdős, P., Rényi, A., 1959. On random graphs. *Publ. Math. Debrecen* 6, 290–297.
- Horgan, G.W., Ball, B.C., 1994. Simulating diffusion in a Boolean model of soil pores. *Eur. J. Soil Sci.* 45, 483–491.
- Klute, A., Dirksen, C., 1986. Hydraulic conductivity and diffusivity: laboratory methods. *ASSA-SSSA. Methods of Soil Analysis, Part 1. Physical and Mineralogical Methods: Agronomy monograph*, no. 9, pp. 687–734.
- Krapivsky, P.L., Redner, S., 2001. Organization of growing random networks. *Phys. Rev. E* 63, 066123.
- Leij, F.J., Ghezzehei, T.A., Or, D., 2002. Analytical models for soil pore-size distribution after tillage. *Soil Sci. Soc. Am. J.* 66, 1104–1114.
- Lin, H.S., McInnes, K.J., Wilding, L.P., Hallmark, C.T., 1999. Effects of soil morphology on hydraulic properties. I. Quantification of soil morphology. *Soil Sci. Soc. Am. J.* 63, 948–954.
- Mooney, S.J., Korošak, D., 2009. Using complex networks to model two- and three-dimensional soil porous architecture. *SSSAJ* 73, 1094–1100.
- Newman, M.E.J., 2003. The structure and function of complex networks. *SIAM Rev.* 45, 167–256.
- Pentier, E., Rieu, M., Sposito, G., de Marsily, G., 1996. Models of the water retention curve for soils with a fractal pore size distribution. *Water Resour. Res.* 32, 3025–3031.
- Price, D.J.S., 1965. Networks of scientific papers. *Science* 149, 510–515.
- Price, D.J.S., 1976. A general theory of bibliometric and other cumulative advantage processes. *J. Am. Soc. Inform. Sci.* 27, 292–306.
- Protz, R., VandenBygaart, A.J., 1998. Towards systematic image analysis in the study of soil micromorphology. *Sci. Soils* 3, 34–44.
- Santiago, A., Benito, R.M., 2007a. Emergence of multiscaling in heterogeneous complex networks. *Int. J. Mod. Phys. C* 18, 1–17.
- Santiago, A., Benito, R.M., 2007b. Connectivity degrees in the threshold preferential attachment model. *Physica A* 387, 2365–2376.
- Santiago, A., Benito, R.M., 2008. An extended formalism for preferential attachment in heterogeneous complex networks. *Europhys. Lett.* 82, 58004.
- Santiago, A., Benito, R.M., 2009. Improved clustering through heterogeneity in preferential attachment networks. *Int. J. Bifurc. Chaos* 19 (3), 1029–1036. doi:10.1142/S0218127409023445.
- Santiago, A., Cárdenas, J.P., Losada, J.C., Benito, R.M., Tarquis, A.M., Borondo, F., 2008. Multiscaling of porous soils as heterogeneous complex networks. *Nonl. Proc. Geophys.* 15, 893–902.
- Strogatz, S.H., 2001. Exploring complex networks. *Nature* 410, 268–276.
- VandenBygaart, A.J., Protz, R., Tomlin, A.D., 1999. Changes in pore structure in a no-till chronosequence of silt loam soils, southern Ontario. *Can. J. Soil Sci.* 79, 149–160.
- Vogel, H.-J., Roth, K., 2001. Quantitative morphology and network representation of soil pore structure. *Adv. Water Resour.* 24, 233–242.
- Watts, D., Strogatz, S., 1998. Collective dynamics of small-world networks. *Nature* 393, 440–442.